STATISTICAL CIRCUIT DESIGN: A SOMEWHAT BIASED SURVEY

by

S. W. Director and L. M. Vidigal

DRC-18-30-81

September 1981

Statistical Circuit Design: A Somewhat Biased Survey[1]

S.W. Director
and
L.M. Vidigal[2]

Department of Electrical Engineering
Carnegie-Mellon University
Pittsburgh, Pa. 15213

Abstract

This paper reviews some of the work in the area of statistical circuit design. While space does not permit a more then superficial survey of the entire field, specific emphasis is placed on those techniques more familiar to the authors.

1 Introduction

Statistical analysis and design of electronic circuits has been an active field of research for more than a decade now. While space does not permit a detailed survey of all methods which have been proposed, (see for example [62,32,8,39]) for good surveys of these methods) we will review some of the more recent work. Particular emphasis will be placed on those techniques that are most familiar to the authors and are specifically applicable to integrated circuit design.

Basically statistical design techniques can be separated into two cfoims. The first encompasses those techniques which are by nature probabilistic, i.e.. the region of acceptability R_{ft} (defined precisely in Section 2) and the optimal nominal design are only approximated in some statistical sense. Typically the nominal design will correspond to the expected value of a probability distribution function and by the process of sampling from this distribution it is possible to make statistical inferences on the location and the characteristics of R_a. The second class of methods are deterministic This class of methods includes those techniques which either try to find a mathematical programming problem whose solution is equivalent to that of the statistical design problem, or else try to approximate the region of acceptability deterministicaily and, based on this approximation, select the best nominal design.

Probabilistic methods require a large number of samples (usually many thousands) to obtain a solution with a reasonable confidence interval. However, since they are essentially independent of the number of designate parameters, one should expect a threshold above which these methods would be computattonatty advantageous. Another advantage of probabilistic methods is their potential capability of convergence to the global solution of the problem independently of any assumptions on the region of acceptability. On the other hand convergence of deterministic method is in general dependent on some convexity assumption as is usual in mathematical programming. Other potential advantages of deterministic methods derive from the fact that the number of independent parameters in integrated circuit design in generally small [45]. Furthermore, the information gained by deterministic methods can be useful to the designer in redefining constraints to ,the problem, and allowing him to gain a deeper insight into the circuit behaviour. Methods such as the simplicial approximation [20], for example, generate an approximation to the region of acceptability which can be used for inexpensive Monte Carlo analysts [21]. Other schemes [63]approximate each constraint. and allow the designer to identify active (relevant) constraints and thereby predict the influence on the yield of lightning or relaxing the values of constraints.

The paper is organized as follows. In the next section we review some of the basic concepts and definition necessary in the sequel. In Section 3 some of the more interesting probabilist used methods for statistical design are desbribed. This should give the reader an idea of the trends of the current work in this area. Section 4 contains a similar summary concerning deternwiistic methods. Finally in section 5 some new results are presented.

2 Mathematical Formulation

In what follows we denote the desfcmabte parameters of a circuit by the n-vector p.. The m circuit attributes which characterize the circuit behavior are called performance functions and are represented as an m-vector f(p). Evaluation of f(p) which is typically of the tern (see for example [19,27,10])

$$f(p) = \int_{t_0}^{t_f} \langle P < x J c 4 M \rangle dt \qquad (1)$$

requires, for nonlinear dynamic networks, the solution of a set of algebraic differential equations

$$N(x_t ip,t) = 0 \quad t_0 \le t \le t_f \qquad (2)$$

where t represents time, and x is a vector of node and branch voltages and currents.Sokition of (2) is refered to as a circuit simulation.

It is convenient to view f(p) as a mapping $fiR^n \to R^m$, where R^n represents the input or parameter s£ac£ and R^m represents the output or performance space. In electronic design problems the performance functions are usually differentiable. We can the vector whose components are the partial derivatives of f the gjarfSTJi of $f(p)_v$ and represent it by Vf(p).

It should be dear from the outset that a function evaluation especially in the form of (2) is a computationally intensive task. The simulation of moderate sized networks can take minutes of CPU time. Therefore, the complexity of any statistical design method will be fundamentally measured by the number of function and gradient evaluations required. The work involved in the backward integration required by the adjoint method [18] for evaluating the gradient has been verified [27] to be of the same order of a circuit simulation.

A particular realization of the circuit corresponding to the parameters p, will be accepted if the performance functions satisfy some upper and lower bounds, i.e., if

$$\hat{} \le KP) < \hat{} \qquad (3)$$

where f^1- and f^0 are m dimensional vectors representing the limits of acceptable perfoimances.We will sometimes express these constraints as $f(p) \le y$. We assume limits on the values p.

[1]This research was supported in part by the National Science Foundation under Grant ECS 79-23191

[2]On leave from institute Superior Tecnfco.Usoon,PQrtugat

$$P^L \leq P \leq P^U \quad (4)$$

White these limits can frequently have a physical interpretation, their main objective is to guarantee the boundedness of the set (5) Constraints (4) will be designated as box constraints. In input space, constraints (3) and (4) define the region oif acceptability:

$$R_a = \{p | f^P - P^P \leq P^U\} \quad (5)$$

which represents the set of points corresponding to acceptable circuits.

We assume that the fluctuations in the manufacturing process are characterized by a unimodai joint density distribution function ^P.Pg)- If the number of circuits fabricated is large, the production yield will be equal to the probability that a circuit outcome will satisfy the specification: i.e.,

$$Y(p^\circ) = \int \ldots \int_R \Phi(p,p^\circ)dp \quad (6)$$

Evaluation of (6) & difficult as it requires the computation of a multiple dimensional integral over a domain, R_a, known only implicitly through equations (2) . Monte Carlo techniques [31,28.60.26] have successfully been used to estimate (6) in [9.65, 3. 52]. However, such methods tend to be very expensive if reasonable confidence intervals to the estimator of Y, are to be obtained.

A different approach to maximize (6) is that of Brayton, *et al.* [11,13], which is basically a generalization of the ideas proposed in [20] for arbitrary statistical distributions. This method consists of approximating the problem of maximizing yield, by a design centering approach which is the geometrical problem of inscribing the largest level set $L_{Q(a)}$ of the distribution $\Phi(p,p^0)$ in the region of acceptability.The lcflrel set $L_o(a)$ is defined as

$$L_{p^o}(\alpha) = \{p | \Phi(p,p^o) \geq \alpha\} \quad (7)$$

The yield maximization problem ,which can be-defined lm

$$YM: \max Y(p^\circ) \quad (8)$$

is therefore replaced by the design centering problem

$$DC: \min_{p^o} \alpha$$

$$s.t \quad L_o(a) \subset R_t \quad (9)$$

the use of this norm has the advantage of making problem DC solvable, as distances can then be related to a norm functional. In general the solutions to YM and DC are different but will in general be dose to one another. If $tp-p°) is unimodai, the contribution to the integral (6) will decrease with the distance to p°. Because (9) attempts to center the distribution in R_a, the region where the values of the integrand function are significant wUJ contribute to the yield. An interesting relation between problems DC and YM is established in [39].

For many practical problems the level set in (9) can be related to a norm n(x). Using this relation it is possible to define a norm fcojjbfc which is the closed convex set

$$B_{p0}(r) = \{p | n(p-p^\circ) \leq r\} \quad (10)$$

Frequently the distribution <I>(p,p°) is gaussian.in which case the norm bodies are ellipsoids and the corresponding norm a weighted second norm. Another interesting case, specially for the design with descrete components is when tolerances are assigned to each parameter. In this case the norm body is a orthotope and the corresponding norm the max or infinite norm. Equation (10) would in this case be written as in (11) where t is a n-vector whose components represent the tolerance assigned to each parameter.

$$B_{p0}(t) = \{y | |y_j - p_j| \leq t_j\} \quad (11)$$

3 Probabilistic Methods

Most methods available for yield estimation are statistical in nature, and mainly consist on variations of the so called "crude" Monte Carlo method [28, 26]. In this method N points $p^1 \ldots p^N$ are chose at random from a population with distribution $\Phi(p, p^\circ)$ and the yield is approximated with the unbiased minimum variance estimator

$$\hat{Y} = 1/N \sum_{j=1}^{N} g(p^j) \quad (12)$$

where

$$g(p) = 1 \text{ if } p \in R_a$$

$$g(p) = 0 \text{ if } p \in R_a \quad (13)$$

The most significant drawback of Monte Carlo method is that to obtain a good confidence interval for the estimator (l 2) a very large value of N must be used as the variance of Y decrease with $1/N^{1/2}$. However, the method is independent of the dimensionality of the multiple integral (6) and of any assumptions on FK These properties make Monte Carlo techniques very attractive for problems with a large number of destgnabie parameters and many researchers have investigated methodologies to adapt Monte Carlo techniques to statistical design problems. Another drawback of Monte Carlo methods results from the statistical character of the information obtained.Given two nominal points p^1 and p^2, if the corresponding yields $Y(p^1)$ and $Y(p^2)$ are estimated by Monte Çarlo, it is difficult to compare the relative merit of.each solution $Y(p^1)$, $Y(p^2)$, particularly if the difference $| Y(p^1) - Y(p^2) |$ is of the order of the confidence interval of the estimators. In other words, gradient information obtained by perturbation is unreliable.

3.1 Importance Sampling

Importance, sampling [28,60] is a scheme to achieve tighter confidence intervals for the estimator of the yield, by modifying the distribution of the population from which the samples are extracted. Notice that (6) can be rewritten as

$$Y(p^o) = \int \ldots \int_Q \frac{\Phi(p,p^o)}{\Phi(p,p^o)} \Phi(p,p^o)\,dp \quad (14)$$

where i is another ftad.f. subjected to the condition that it can be zero only when $\Phi(p,p^\circ)$ is zero. Sy drawing the N samples from the distribution f the yield can be estimated as

$$Y = 1/N \sum_{i} g(p^i) \frac{\Phi(p^i,p^o)}{\Phi(p^i,p^o)} \quad (15)$$

the variance is now giveaby

$$Var = 1/N \int \ldots \int [g(p) - Y]^2 \Phi(p,p^o)\,dp$$

this variance is minimized, actually made equal to zero if we select

$$r(p,p^\circ) = \frac{\Phi(p,p^o)}{\int \ldots \int \Phi(p,p^o)\,dp} = \frac{\Phi(p,p^o)}{Y(p^o)}$$

Of course to obtain such a distribution Y(p°) has to be known and this dearly it is not the case. However if the designer has some knowledge of the problem a distributions can be devised that greatly reduce the variance of the estimator. Importance sampling techniques under various forms have been proposed for statistical design [21, a 52].

3.2 The Method of Antreich and Koblltz

An interesting application of equation (14) was suggested recently by Antreich and Kobiitz [1]. Given a nominal design p°, maximizing yield corresponds to finding a new nominal point .say p^1 and a new distribution (p,p^1) such that the integral (6) will be maximized. If in (14) we make $F(p,p^\circ) = \Phi(p.p^1)$ the yield afp^1 is given by

$$Y(p^1) = \int \ldots \int_{R_a} \frac{\Phi(p,p^1)}{\Phi(p,p^o)} \Phi(p,p^o)\,dp$$

$$= \int \ldots \int_{R_a} g(p) \Phi(p,p^o)\,dp$$

The yield at p^1 is then the expected value

$$Y(p^1) = E.V.\left[g(p) \frac{\Phi(p)}{\Phi(R,P)}\right] \quad (16)$$

Assuming that the distributions are normal with the same covariance matrices. (16) is expanded in series around p° to obtain

a quadratic function in terms of the difference p^1-p°. The parameters required for the expansion can be evaluated by means of a Monte Carlo analysis around p°. This quadratic can then be maximized to obtain the ideal value of p^1 although some care must be taken as the approximation is local and the quadratic is not in most cases concave.

3.3 The Method of Bias [22]

Observe that the integral (6) can be extended over the whole space in the following way

$$Y(p^\circ) = \int_{R_a} g(p)\ \Phi(p,p^\circ)\ dp \qquad (17)$$

where g(p) was defined in (13). Furthermore the gradient with respect to p° can be found by differentiating (17), and noting that g(p) is not a function of p°

$$\nabla Y(p^\circ) = \int_{R_a} g(p)\ \nabla\Phi(p,p^\circ)\ dp = \int_{R_a} g(p)\ \frac{\nabla\Phi(p,p^\circ)}{\Phi(p,p^\circ)}\Phi(p,p^\circ)\ dp = \int_{R_a} g(p)\ \nabla \ln \Phi(p,p^\circ)\ \Phi(p,p^\circ)\ dp \qquad (18)$$

Hence from equations (17) and (18) we see that the yield and its gradient are the expected values:

$$Y(p^\circ) = E.V.[g(p)]$$

$$\nabla Y(p^\circ) = E.V.[g(p)\nabla \ln \Phi(p,p^\circ)] \qquad (19)$$

If the N samples are generated as before, then the yield can be estimated as in (12) and .simultaneously , we can estimate the gradient as

$$\nabla Y(p^\circ) \approx \frac{1}{N} \sum_{i=1}^{N} g(p^i)\ \nabla \ln \Phi(p^i,p^\circ) \qquad (20)$$

Thus the amount of additional work to compute the gradient is very small as closed expressions for $\nabla \ln \Phi(p, p^\circ)$ are normally available. However given this information we can consider using it to adjust the nominal so as to improve yield. A similar idea has been developed by Bias in [22,23] . Although computationally very intensive, the proposed method represents a significant improvement over early techniques which were basically random searches for the maximum yield [38,9]. In Bias method, rather than use a direct formula as (20) to determine gradients he introduces the following expression to measure the sensitivity of the yield with respect to a designate parameter

$$S_i = \frac{P_{pass}(p^i) - P_{fail}(p^i)}{P_{pass} + P_{fail}} \qquad (21)$$

where P_{pass} and P_{fail} represent conditional probability densities for p^1 given that for the value p^1 the circuit respectively passed or failed the specifications. The extreme values of M, range from zero for a totally insensitive parameter to two for a parameter which completely controls circuit yield. The idea is to rank parameters in order of sensitivity and assign new tolerances to the more sensitive parameters by assigning the range of values for which $p_{pass}(p^i) \geq p_{fail}(p^i)$.

3.4 The Method of Soin *et al* [61, 30]

This is a conceptually very simple approach to design centering, which can be expected to produce significant increases in the yield in most applications. The method [61,30] is also based on repetitive application of Monte Carlo analysis. For each analysis the center of gravity of the samples lying outside and inside the region of acceptabiiity,denoted by g° and g\ respectively, is used to generate a direction of search, and the new nominal is chosen as

$$p^N = p^\circ + \lambda (g^i - g^o) \qquad (22)$$

where \ is a step size normally taken to be one.

The concept can be better illustrated by refering to figure 1, where the results after drawing 5 samples (p^1, p^2, ..., p^5) are illustrated.

After 5 samples a new nominal design p° is found,such that the tolerance box is a subset of R_a and hence the corresponding yield is 100%. However the method would probably have difficulties in converging to acceptable solutions in some other situations.

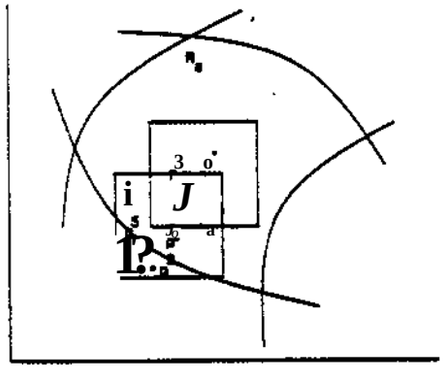

Figure 1: illustration of Soin method

3.5 The Method of Kjellstrom el *al* [35, 36]

This method could be seen as a generalization of method described in the previous section and is probably the most promising technique for stochastic design. The method [33,34,37,35,36] consists of exploring the input space with an adaptative gaussian distribution which keeps the probability of hitting the region of acceptability essentially equal to 1/e » 0.36a Motivation for this method is based on the fact that from all distributions with equal second order moments, the gaussian distribution has the largest dispersion and hence is the one which best covers the input space.Furthermore, the best rate of information is obtained if the probability of getting a sample m R_a is kept at 1/e.

Kjellstrom seeks to characterize the region of acceptability by its center p* (first order moment) and extension M_a, (second order moment), given by

$$\mu^a = \frac{1}{Y}\int_{R_a} p\ \Phi_N(p,\mu^a)\ dp$$

$$M_a = [m_a(i,j)],\ m_a(i,j) = \frac{1}{Y}\int_{R_a} p_i p_j\ \Phi_N(p,\mu^a)\ dp \qquad (23)$$

where

$$Y = \int_R \Phi_N(p,\mu^a)\ dP \qquad (24)$$

and $ (p,p) is afgaussian distribution with mean p* and correlation matrix M_t, which is updated after each success to keep Y approximately constant and equal to 1/e.

4 Deterministic Methods

This class is characterized by methods which solve the statistical design problem by mathematical programming techniques, and has followed closely the progress made in this area,

4.1 Formulation as a Constrained Minimization Problem

A yield of 100% can be obtained by assigning tolerances to each component In this situation the norm body is defined as in (11) and a natural formulation of the 100% yield problem would therefore be

$$\min\ c(p^\circ,t)$$

$$s.t\ \ B_p o(t)\ C\ R_a \qquad (25)$$

where c(p°, t) is some cost function reflecting the prices associated with each element and its tolerances.A typical cost function is

$$c = \sum_{k=1}^{n} \frac{t_k}{t_k} \qquad (26)$$

where p° represents the nominal design and t is a vector representing tolerances. In most applications [50,51.4,5] only a few of the infinite number of constraints in (25) are used. These constraints normally correspond to a subset of the 2" vertices of the tolerance orthotope, which by some experimentations the designer expects to be the critical ones.

An equivalent formulation which was recently proposed by Poiak and Sangiovanni-VincenteHi [53,54] expresses (25) as[1]

$$\min_{p^0,t} c(p^\circ,t) \qquad (27)$$

$$\text{s.t} \min_i \max_\mu f_i(p) \le y \;,\; p \text{ » } p^\circ \bullet t p$$

where ft is a vector whose elements are -1 ≤ \L ≤ 1. Polak and Sangiovanní-VincenteHi showed that the constraints in (27) are Upschitz and thus nondifferentiable. They suggest algorithms to handle this type of problem. Their formulation is very general as it can be extended to include tuning and is not dependent on assumptions of convexity of the region of acceptability. Convergence is established to a locally optimal solution of (27), however no results have been published and the efficiency of the algorithm is difficult to judge. The area of optimization of nondifferentiable functions is an active one and significant improvements might be expected in the near future.

4.2 The Method of Madsen and Schjaer-Jacobsen [43, 59]

Madsen and Schjaer-Jacobsen [44,59] have extended their early work in minimax optimization to the problem of statistical design. Their method represents a very natural formulation of toierancing problems, which is easy to implement and has the desirable feature of quadratic convergence in many applications. In particular, their method is geared towards solving the following minimax problem [40,41,42,43]

$$\min_P \max_i f_p(p) \qquad i \text{ » } 1,2,..M \; m \qquad (28)$$

An iterative procedure is used, which at every iteration replaces the functions by their Taylor expansion:

$$\min_{p^{k+1}} \max_i f_j(P_k).+ \nabla f_i(p^k)^T (p^{k+1} \cdot p^k)$$

$$\| p^{k+1} \cdot p^k \|_\infty \le \lambda_k \qquad (29)$$

where \ is a bound on the maximum step allowed. This bound is necessary to prevent (29) from having an infimum at infinite. The parameter \ is adjusted automatically at each step according to the goodness of the linear approximation and will be small near the solution. This update procedure is empirical and further details can be found in [40]. The minimax problem (29) can be solved efficiently as a linear program, by introducing an additional variable γ:

$$\min \; 7$$

$$\begin{cases} f_i(p^k) + \nabla f_i \text{«} p^{kT}) h^k \le 7 & i = 1, \ldots, m \\ h_j^k \le \lambda_k & j = 1, \ldots, n \\ \bullet h_j^k \le X_k & i = 1, \ldots, n \end{cases} \qquad (30)$$

where h^k » $p^{k+1} \bullet p^k$. If m>n, it is very often the case that dose to the solution n+1 constraints will be active, this means that the minimum of the function

$$F(p) \text{ » } \max_i f_i(p) \qquad (31)$$

is attained at a point where F(p) is not differentiate. In this situation the minimax problem is designated as rgflyjai and it can be established that in this case the rate of convergence of the algorithm is quadratic [43]. If at the solution n+1 constraints are not active, the problem is called singular, in which case the convergence of (30) depends critically on the value of \ and therefore will normally be slow.

To extend the minimax concept to statistical design problems, Madsen considers three sequential design steps in increased order of complexity. The first, which is referred to as the Zero Tolerance Problem, corresponds to the solution of (28) and is basically a multiple objective optimization problem. The second step corresponds to the Eixed Tolerance Problem case in which the tolerances are fixed and the algorithm seeks a nominal design such that the maximum possible value of the performances is minimized. The third step, referred to as the Variable Tcierance Problem is the most complex one where both tolerances and nominal are to be designed. The approach used for this problem is quite inefficient as a Regula-Falsi technique is applied to a sequence of Fixed Tolerance Problems.

The algorithms can be summarized as:

1. The worst case problems (WCP), may be stated formally as

$$\text{WCP:} \quad \text{<Pi(p,r)} \text{ » } \max_{y \in a_r(p)} f;(y)i \text{»} i,2.\ldots.m \qquad (32)$$

and corresponds to finding in the tolerance orthotope centered at p, a point where the function f_i assumes its maximum value.

2. The fixed tolerance oroblemfFTPl is formulated as in (33) and is of course a function of the tolerances.

$$\text{FTP:} \quad 4\text{>(r)amin max <q.(p,r)} \qquad (33)$$

3. Note that the limiting case where r^o, known as the jerfl tolerance problem (ZTP) and is basically an optimization problem without toierancing

$$\text{ZTP:} \quad 0(0) \text{ «i} \min_P \max_i f.(p)$$

4. Finally the variable tolerance problem can be stated as (34)

$$\text{VTP:} \quad \Phi(r) = f_{max}$$

or

$$\min_P \max_i \max_{y \in B_f(p)} \text{':(y>«<}_{max} \qquad (35)$$

for a given value of f_{max}

Note that the problems FTP, ZTP and VTP are in a form identical to (28) and can therefore be solved using the algorithm described above. Notice also that FTP and VTP require the solution of a WCP and this subprobiem has to be solved at every iteration of the algorithm. Because this can be a nontrivial optimization problem a simplifying assumption is made that the maximum of thei, occur at the vertices of the tolerance polytope.

4.3 The Simplictal Approximation Approach

Director and Hachtei [20] proposed a technique for design centering which represents a significant departure from previous techniques. The method is an iterative procedure which generates points on the n-1 dimensional surface R_a. Assuming convexity at each iteration, an interior approximation to R_a is obtained. This approximation is the polytope $?_k$ - co{v\ v^2. . . . v^{11}} whose vertices are the points on the boundary of R_a. P_k can be characterized by

$$P_k\text{-}\{P\,l/p\,\hat{}\,b_{.,}, \; i\text{-}1,Z\ldots,n_{p_k}\} \qquad (36)$$

where the c* represents the outward pointing normals to the faces of the polytope, the b_i the distance of the face to the origin and n_F the total number of the faces at iteration k. Normally the initial polytope P_Q is a simplex which is the convex hull of n • 1 initial points on the surface of R_a, which are obtained by line searches in the coordinate directions. At every iteration the largest normbody (10) is inscribed in the polyhedral approximation P_k.Because the distance from a point p to a hyperpiane is a linear function of p. this inscription can be carried out efficiently with the following linear program!

$$\begin{aligned} &\max \; r \\ &\text{it} \quad c^* \; P + r n\hat{}\hat{}bj \;,\; j\text{«}1,2\ldots., n_{p_k} \end{aligned} \qquad (37)$$

At each iteration one of the active constraints in (37), or equivalents, one of the faces touching the largest inscribed

[1]The multiplication t *ft* is assumed to be term by term

normbody is broken and a search made to locate a new point on the boundary of R_a. This point, v^{k+1}, is added to the set of vertices of P_k and the new approximation to R_a is obtained:

$$P_{k+1} = co\{P_k \cup v^{k+1}\} \tag{38}$$

The procedure is terminated when there is a small variation of center of the inscribed normbody for two consecutive iterations.

In the original paper [20] the new vertex of the polytope was determined by breaking the largest of the faces of the polytope. The size was determined by measuring the radius of the largest n-1 sphere inscribed in the active face. A line search is then made outwards from this center to obtain a new vertex of the polytope. Many other schemes are possible for updating the polytope. Sangiovanni-Vincentelli [64] proposed four different update rules and proved that for any of these updates the algorithm would converge to an optimal solution of the design centering problem.

4.4 The Method of Brayton *et al* [12]

The approach followed is similar to the proposed by Madsen and Schjaer-Jacobsen [59, 41] which was described above. Since in this method only first order information is used to solve the FTP, a poor convergence rate usually results when the minimax problem is singular. The method proposed by Brayton *et al* [12] to solve the FTP and the VTP differs in two main aspects. First, the FTP is transformed into a constrained mathematical programming problem and solved with a quasi-Newton algorithm. This allows for second order information to be obtained and better global and local convergence should be expected. The second difference concerns the approach to the VTP, which is also converted into a constrained problem and solved directl, as explained below.

The ZTP can be transformed into a constrained optimization problem in the following way:

$$\min_{p,\gamma \in R^{n+1}} \gamma$$

$$\text{s.t. } f_i(p) - \gamma \leq 0 \quad i = 1,2,\ldots,m \tag{39}$$

An efficient way of solving (39) is the use of a variable metric algorithm for constrained optimization. The method suggested is a variation of the algorithms proposed by Han [29] and Powell [56, 57, 58] (Han-Powell algorithm). Notice that a straightforward application of such an algorithm to (39) would not take into account the special structure of the problem, i.e. that it is linear in γ. For this reason the Han-Powell algorithm is modified in the following way. At each iteration an approximation to the solution p^k, γ is known as well as a set of Lagrange multiplier λ^k. The Han-Powell algorithm would obtain the next direction of search, by solving the following program.

$$\min_{d,\delta} \frac{1}{2} (d,\delta)^T Q_k (d,\delta) + \delta$$

$$\text{s.t. } f_i(x^k) + \nabla f_i(x^k)^T d - \gamma - \delta \leq 0 \quad i = 1,\ldots,m \tag{40}$$

where Q_k is a positive definite approximation to the Hessian with respect to p of the Lagrangian function of (39).

$$L(p,\gamma,\lambda) = \gamma + \sum_{i=1}^{m} \lambda_i (f_i(p) - \gamma) \tag{41}$$

Due to the special structure of the problem, linearity in γ, the matrix

$$\nabla^2_{p,\gamma} L = \sum_{i=1}^{m} \lambda_i \nabla^2_{p,\gamma} f_i(p) \tag{42}$$

has only zeroes in the last row and column. Therefore, the approximation to the inverse of $\nabla^2_{p,\gamma} L$ should be initiated and kept as:

$$Q_k = \begin{bmatrix} \tilde{Q}_k & 0 \\ 0 & 0 \end{bmatrix} \tag{43}$$

and the quadratic program (40), has the simplified form

$$\min_{d,\delta} \frac{1}{2} d^T \tilde{Q}_k d + \delta$$

$$\text{s.t. } f_i(x^k) + \nabla f_i(x^k)^T d - \gamma^k - \delta \leq 0 \quad i = 1,2,\ldots,m \tag{44}$$

The FTP is solved taking a similar approach. The equivalent form for this problem is

$$\min_{p,\gamma \in R^{n+1}} \gamma$$

$$\text{s.t. } \varphi_i(p) - \gamma \leq 0 \quad i = 1,2,\ldots,m \tag{45}$$

where $\varphi_i(p)$ is defined in (32) and represents the solution of a WCP centered at p with a fixed value or r. Problems (45) and (39) have a similar formulation. but solving (45) is a much harder task for two reasons. Evaluating $\varphi_i(p)$ requires the solution of a WCP, and this subproblem may require a significant number of gradient evaluations. Also the functions $\varphi_i(p)$ are not differentiable and therefore the class of algorithms available to solve (45) is smaller than for (39).

Assuming the vertex hypothesis,i.e.,that the maxima of f_i will occur at the vertices of the tolerance orthotope, it is possible to convert (45) into a constrained optimization problem with differentiable constraints as follows. Let v(j) represent the jth vertex of the tolerance orthotope and define $f_{ij}(p) = f_i(p + r \times t \times v(j))$, i.e. the value of function f_i at vertex j of the tolerance box $B_r(p)$. Under the aforementioned hypothesis, inequalities (46) and (47) are equivalent.

$$\varphi_i(p) = \max_{y \in B_r(p)} f_i(y) < \gamma \tag{46}$$

$$f_{ij}(p) \leq \gamma \quad j = 1,2,\ldots,2^n \tag{47}$$

Problem (45) would, under these assumptions, be equivalent to:

$$\min_{p,\gamma \in R^{n+1}} \gamma$$

$$\text{s.t. } f_{ij}(p) = f_i(p + r \times t \times v(j)) \leq \gamma \quad \begin{matrix} i = 1,2,\ldots,m \\ j = 1,2,\ldots,2^n \end{matrix} \tag{48}$$

notice that there are $m \times 2^n$ constraints, and solving a nonlinear problem with such a dependence on the number of designable parameter soon becomes very difficult or impossible.[2] It is desirable to keep the number of constraints reasonably bounded. This could be achieved by keeping only the f_{ij} corresponding to the vertices which are solutions of the WCP, but the global solution of this problem is difficult to obtain and may occur at more than one vertex. To overcome this difficulty,a list L_i is kept of all vertices which are probable candidates to be the maximum of f_i over the tolerance polytope. Then in (48) only the indices j existing in L_i are considered to obtain (49). As a result only a moderate number of constraints is kept. These constraints can be updated dynamically (with L_i) so as to always include the vertices of $B_r(p)$ which are good candidates to correspond to active constraints in (48). The method used to update L_i is described below.

$$\min \gamma$$

$$\text{s.t. } f_{ij}(p) \leq \gamma \quad i = 1,\ldots,m \quad j \in L_i \tag{49}$$

The algorithm to solve (49) is similar to the one described for the fixed tolerance problem, except that before solving the quadratic program (40) a worst case algorithm is applied to each performance function. One iteration of the algorithm can be stated succintly as follows:

1. Solve the worst case problem for each function $f_i, i = 1,\ldots,m$.
2. Add to the vertex list L_i all vertices visited during (1). Note: function and gradient evaluation are expensive,

[2] We notice that (48) is equivalent to the formulation used in [50, 51, 4, 6]

and it is natural to use ail obtained information in step (3).

3. Solve the quadratic program, equivalent to (44)

$$\min_{d,\delta} \tfrac{1}{2} d^T Q_k d + \delta$$

$$f_{ij}(p^k) + \nabla f_{ij}(p^k)^T d - /.5 < 0 \quad i = 1....m$$

$$j \in L_i \qquad (50)$$

where

$$f_{ij}(p) = f_i(p + r\,v(j))$$

$$\nabla f_{ij}(p) = \nabla f_i(p + r\,v(j))$$

4. Drop from the vertex list L_i all vertices which did not correspond to active constraints in (50), except if that was the vertex corresponding to the global maximum in (1).

5. Check the step (d,5) can be accepted. If not reduce the step size.

6. If a step is rejected, add to the vertex fist the number associated to the vertex that caused the rejection. When the step is accepted add to the vertex list all vertices visited during the WCP.

7. Go to a

To solve VTP, (35) is reformulated as

$$\max r$$

$$\& \quad \wedge(P.O - f\wedge\wedge O \quad i\text{-}IA\text{-}Mfli \qquad (51)$$

it is important to note the difference between this problem and problem (45). In (45), r is fixed and the algorithm win try to situate the constant sized tolerance box in a position where the maximum possible function value is minimized. In (51) the maximum permissible value is given and the objective is to select a nominal design and the maximum value of r such that the constraint $f\wedge$, is not violated.

The procedure used to solve (51) is essentiafly identical to the one described in the previous section. The major difference concerns the structure of the Hessian, which does not have, as before, a column and a row of zeros. The gradient of the constraints with respect to r are, in general, not constant as in (39).

4.5 The method of Bandler and Abdel-Malefc (71

if the region of acceptability is not convex, polyhedral approximations become a poor representation to R^ as win be further explained in the next section. A possible method to overcome this problem is to appoximate the performance functions with higher order polynomial functions. Such a method was first proposed by Bandler and Abdel-Malek [7], who for pratical reasons limited the polynomials to be of second order. The performance functions f_i are initially approximated by quadratic polynomials:

$$P_i = a_1 p_1^2 + \ldots + a_{N-1} p_{N-1} + a_N \qquad (52)$$

where the $N = (n+i)(n+2)/2$ coefficients are determined by evaluating the function value at N different points, normally around the initial nominal design. These approximations can then be used as a cheap substitute of the performance functions,to solve an optimization problem such as for example (27), because evaluating the value and the gradients of (52) is a trivial task. Note that if the points chosen initially to evaluate $P_i(p)$ are not close to the active vertices of the tolerance orthotope , then the quadratic approximations can be a poor representation of the constraints. New updates to the polynomials must then be recomputed and the optimization step repeated. This scheme should be compared with the method proposed in the next section where an effort is made to approximate locally the constraints in the regions where its relevance to the inscription problem is expected.

An algorithm for optimal tolerance assignment for yields smaller than 100% is also proposed. The following nonlinear program is set up:

$$\text{minimize} \quad c(p^\circ, t) \qquad (53)$$

$$\text{subject to} \quad Y(p^\circ, t) \geq Y_L$$

where $Y(p^\circ,t)$ represents the yield for a nominal p° and a tolerance vector t . Y_L is the minimum acceptable yield, a number specified by the designer. To solve (53) the yield and its gradient must be evaluated. The technique used makes a number of further assumptions, besides the use of the quadratic approximations described above. First the intersections of the approximations with the edges of the truncated p.d.f. (tolerance polytope) are used to further linearize the quadratic approximations. Linear cuts of the tolerance orthotope are thus obtained . If the p.d.f. is uniform and the cuts do not overlap, than the yield is a simple ratio of volumes. However for general distributions the evaluation of the yield is difficult One approach, in this case, is to regionalize the tolerance orthotope into smaller orthotopic cells. A Monte Carlo analysis is then carried to estimate the probability of the parameters falling in each of the cells.

5 A Second Order Approximation to R_a

5.1 Introduction

The Simpliciai Approximation, method for design centering, which we described in Section (4.3) was developed under the assumption that R_a is convex. When R_a is not convex, the solution obtained using the Simpliciai Approximation algorithm can be deceptive. Fig 2 illustrates a hypothetical two dimensional situation where two constraints define a nonconvex region.

Assuming that the level sets associated with the joint density functions are 2-spheres, it would be desirable to determine a solution dose to c^2. However, in this hypothetical case the Simpliciai Approximation algorithm would converge to a solution dose to c^1 and not only result in an infeasible nominal design, but also in a very small yield.

In a typical engineering design situation one should not expect a convex region of acceptability, examination of many examples [14,5,49.25]). seem to indicate that the region of acceptability can be viewed as the intersection of sets which are either cojgg or ttMTiotementary convex fi.e.. a set WTK)S6 cwnptement is convex) [2]. The most general dass of performance functions which result in feasible regions which are intersections of sets which are convex and complementary convex are those that are *ouasL* convex or quasi-concave. Thus we are motivated to consider this dass of functions in some detail.

Definition 1: A function $hR^n \to R^1$ is Quasi-convex if given p^1, $p^2 1 R^\circ$, then for any value of the scalar $, $0 \leq \theta \leq 1$

$$f(\theta p^1 + (1-\theta)p^2) \leq \max [f(p^1), f(p^2)] \qquad (54)$$

A function f is ouaai-concave if its negative, -f, is quasi-convex.| •

If the function is dtfferentiable it can be shown [46] that (54) can be replaced by the following implication:

$$f(p^2) \leq f(p^1) \Rightarrow \nabla f(p^1)^T (p^2 - p^1) \leq 0 \qquad (55)$$

i.e., if $fCp^2) \leq f(p^1)$ then the directional derivative at p^1 is nonincreasing.

Our interest in quasi-convex functions derives from the following property, [46]

Theorem 2: Let f be a real value function defined on a convex set. Then f is quasi-convex if and only if the set

$$L_{fl} * \{p | f(p) \leq a\} \qquad (56)$$

is convex for each $\alpha \in R$.

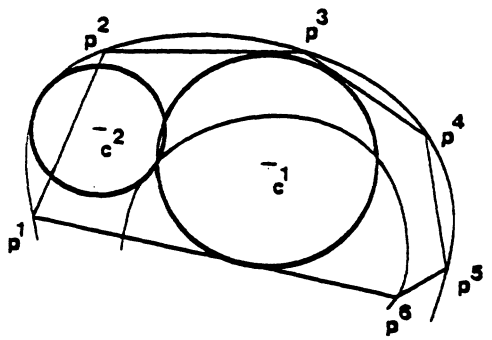

Figure 2: Hypothetical nonconvex region of acceptability. SA algorithm would produce as a solution c^1. Correct solution c^2.

We will often refer to the set L_α as defined in (56) as a level contour of the function f. For the class of quasi-convex and quasi-concave functions the constraints (3) imply that the region of acceptability will be the intersection of the convex sets C_j and complementary convex sets D_j where

$$C_j = \{p | f_j(p) \le f_j^U\} \quad \text{if } f_j \text{ is quasi-convex} \qquad (57)$$
$$= \{p | f_j(p) \ge f_j^L\} \quad \text{if } f_k \text{ is quasi-concave}$$

and

$$D_j = \{p | f_j(p) \ge f_j^L\} \quad \text{if } f_j \text{ is quasi-convex}$$
$$= \{p | f_j(p) \le f_j^U\} \quad \text{if } f_j \text{ is quasi-concave} \qquad (58)$$

to ease the notational burden we define two new functions $g_j(p)$ and $h_j(p)$ in such a way that

$$C_j = \{p | g_j(p) \le g_j^U\} \quad j \in J_g = \{1,2,...,n_c\}$$
$$D_j = \{p | h_j(p) \le h_j^U\} \quad j \in J_h = \{1,2,...,n_c\} \qquad (59)$$

where the g_j are quasi-convex functions and the $h_j(p)$ are quasi-concave functions, which we will for convenience refer to, respectively, as convex constraints and complementary convex constraints.

Note that g_j and h_j are both in absolute value equal to f_j. For the description of the algorithm, separation of the constraints into two types is very helpful, but it should be noted that in a real design problem the functions f_j are only known implicitly through a system of nonlinear differential equations and therefore defining the g_j and the h_j can in itself be a nontrivial problem.

In order to have the boundaries well defined we make a final assumption about the performance functions. Specifically we assume that the gradients of f_j will not vanish at the boundary of R_a.

5.2 Main Algorithm

If the region of acceptability is convex convergent methods such as the simplicial approximation algorithm [20, 64] exist for solving the DC problem. We now show that a stationary solution to DC in the more general case can be found by solving a sequence of problems where the region of acceptability is replaced by a convex approximation.

If the complementary convex constraints are replaced by supporting hyperplanes, the resulting approximation to R_a, $\bar{R}_a$ is convex. This situation results because $\bar{R}_a$ is the intersection of the convex sets C_j with half spaces, and approximates R_a interiorly.

For a given nominal point c, the largest norm body that can possibly be inscribed in R_a is limited by the points on the boundary of R_a which are closest to c. We are assuming, of course, that the measure used for the distance is the norm associated with the distribution $\Phi(p,c)$ as mentioned in section 2.

These considerations suggest the following algorithm:

Main algorithm

Step 0 Let c^0 be an acceptable point, i.e., $c^0 \in R_a$

Let $k = 0$, $J = J_h$

Step 1 For all $j \in J$ solve the following near point problems

$$NP_{k,j}: \min \ n(p-c^k)$$
$$\text{s.t.} \ g_i(p) \le g_i^U \quad i = 1,2,...,n_c$$
$$h_j(p) = h_j^L$$
$$h_k(p) \le h_k^U \quad k = 1,...,j-1,j+1,...,n_c \qquad (60)$$

Step 2 a) If $NP_{k,j}$ has no solution, drop constraint j from further consideration, i.e. $J = J-j$

b) Else, if $p^{k,j}$ is a solution to $NP_{k,j}$, evaluate the approximations

$$\bar{h}_{k,j}(p) = \nabla h_j(p^{k,j})^T(p-p^{k,j}) + h_j^L$$

Step 3 Solve the yield maximization problem, with the convex region of acceptability

$$YMC_k: \max_{c,r} r$$
$$\text{s.t.} \ g_i(c + r\omega) \le g_i^U \quad i \in J_g$$
$$\bar{h}_{k,j}(c + r\omega) \le h_j^U \quad j \in J \qquad (61)$$

Step 4 If the solution to YMC_k is locally optimal to DC, then STOP

Step 5 Go to Step 1

One step of the above algorithm is illustrated in Figure 3 where the norm body is assumed to be a circle.

Given the center of the last inscribed circle, c^k, we solve subproblems $NP_{k,1}$ and $NP_{k,2}$ to find the points $p^{k,1}$ and $p^{k,2}$ on the surfaces $h_1(p) = h_1^L$ and $h_2(p) = h_2^L$. The convex approximation to R_a is then built by introducing the two half spaces $\bar{h}_1(p) \le h_1^U$ and $\bar{h}_2(p) \le h_2^U$. The new center c^{k+1} is then determined by inscribing the largest circle in the convex region.

Remarks

1. We assume the knowledge of an initial feasible design c^0, which exists in the interior of R_a.

2. If no solution can be found to $NP_{k,j}$, then the constraint h_j is superfluous in the sense that it does not constrain the design, and can be dropped, giving useful information to the designer.

3. Step 3 assumes the existence of a convergent method for the case where R_a is convex. We recall that such methods exist, e.g. the simplicial approximation algorithm.

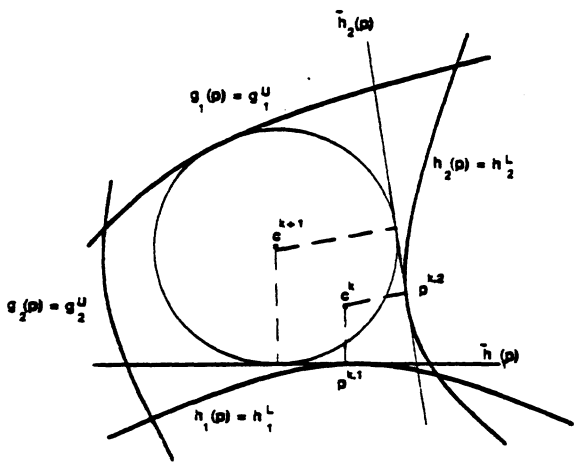

Figure 3: Illustration of a step of Main Algorithm

It is possible to prove that such an algorithm will converge to a solution of the problem DC defined in (9) [63].

Notice that as stated here, the algorithm is not immediately implementable because for a general problem there is no previous knowledge of which constraints are convex and which are complementary convex. In the next section we will further discuss this problem and suggest practical ways to resolve this difficulty, as well as methods to reduce the computational effort associated with the use of the algorithm.

5.3 Implementabfe Algorithm

The algorithm proposed in the previous section is a significant tool to solve design centering problems for nonconvex regions of acceptability. There are however two main points which do not make the method computationally attractive. The first concerns the solution of subproblem YMC_k. as most methods available for solving this problem require a large number of possibly expensive circuit simulations. Furthermore, if the problem has to be solved repeatedly, even if information is shared in successive iterations, the computational requirements will make the method unattractive, particularly for problems in the time domain. The second difficulty concerns the identification of the quasi-convex or quasi-concave characteristics of the performance functions. These functions are known only implicitly through the system of equations that describes the networks behavior. Hence, using a criterion fiks (54) would require the verification of an infinite number of inequalities. Methods to finitely test quadratic functions are known, however, and will be explained below.

Significant computational savings can obviously be obtained if we decide to approximate the constraints by, for example, the first three terms of a Taylor expansion. Global approximations to the constraints are not in our opinion as attractive as local ones. From the main algorithm introduced in the previous section it can be infered that ideally we would like to approximate the cunsliafrits at those points in every constraint which minimize the distance to the initial design center c°. If this design is a reasonable one, the final solution is not in general very distant although significant increases in yield can be obtained by the process of centering. If this is not the case the designer could always re-evaluate the approximation.

The flow of the Implementable Algorithm « summarized in the following steps:

SteftQLetc°eR*, k«0, J-CiA-.-n^

Steo i£ Solve the following near point problems for $i = 1,\ldots,n_c$

$$NP_{UJ}: \quad \min \; n<p-c^{\circ}) \quad \text{s.t} \quad {}_{,}I(p)^{*}\, f_i^{UJ} \quad f_i(p) \geq f_i^{L} \quad f_i^{L} \leq f_i(p) \leq f_i^{L} \quad i \neq I \tag{62}$$

$$NP_{\llcorner i}: \quad \min \; n(p-c^{\circ}) \quad \text{s.t.} \quad f_i(p) \leq ff \quad f_i(p) = f_1^{L} \quad \wedge \; (p) \quad i^{*}\!\!« \tag{63}$$

Step IB

1. If there is no solution to one of the near point problems drop the constraint

2, Else, evaluate the gradient Vf.f) and the Hessian $V^2 f_1(\bullet)$ at the solutions p^{UJ} and p^{u} of each problem

Step i£ From the Hessian and the gradient infer the convexity characteristics of the constraints

Steo 12. Get local quadratic approximations to the constraints e.g.

$$q_i^{U} = f_i^{U} + \nabla f_i(p^{U,i})^{T}(p-p^{U,i}) + \frac{1}{2}(p-p^{U,i})^{T}\nabla^2 f_i(p^{U,i}\text{xp.pu.ij} \tag{64}$$

and with the information of 1C, rename it as g, or h_i

Go to Step 2

§S2Qi£ Solve the near point problem as in Step 1 of the main algorithm

Steo *g* Linearize the complementary convex constraints as in Step 2 of the main algorithm

Steo 3 Solve the yield maximization problem YMC_k

SIfifl 4 If the solution is locally optimal for the quadratic approximations, then stop

Step 5 Go to Step 1E

Remark 1

Steps 1E to 5 correspond to the steps in the main algorithm proposed above, except for the substitution of the constraints by their quadratic approximations. Steps 1A to 10 are necessary to evaluate the approximations and to discriminate quasiconvex and quasiconcave constraints.

Remark 2

Evaluation of the Hessians can be done efficiently, if the optimization algorithm used for Step 1A uses gradient information. We use the Han-Powell algorithm [56,57,58] for constrained optimization is used, and the gradients obtained at each step are employed to build an approximation to the Hessian with an update formula. However, two important differences must be accounted for, in what concerns selecting an update formula. The first difference concerns the direction of the step, which is not defined by the update but solely by the optimization algorithm. The update should therefore be capable of handling arbitrary directions. It is weH known that the symmetric rank one update:

$$Q^{*} = Q + \frac{(y-Qs)(y-Qs)^{T}}{s^{T}(y-Qs)} \tag{65}$$

$S» p^{*}-p \qquad y» Vf(p^{*})-Vf(p)$

only requires that the dtrecions of search be linearly independent [17]. If the new direction to be used is not independent from the previous ones, and the function is quadratic then the denominator of (65) wiU vanish [55] making the test particularly easy. The second difference results from the fact that for most optimization procedures the approximation to the Hessian has to be kept positive definite. In our case we want the update to approximate as closely as possible the Hessian even if it contains negative eigenvaiues.Update (63) has also this desirable property. .

Remark 3

Characterization of a general quasi-convex function isa problem which is still not well solved [16]. Therefore we replace the problem of characterizing quasiconvexity of the performance function by a similar identification of the quadratic approximation:

$$q(x^2) \leq q(x^1) \text{ -» } \nabla q(x^1)^{T}(x^2 - x^1) \leq 0$$

$$\sim (Q x^1 \bullet g)^{T}(x^2 . x^1) \leq 0 \tag{66}$$

Although (66) does not allow the property of quasiconvexity to be finitely tested immediatty, this is possible by checking the eigenvalue of amatrix related to the Hessian (see [47,48.15,24]).

Summary

In this paper we have reviewed some of the recent work in the area of statrirKT?:* design. Space has not permitted us to cover all approaches in equal depth. However the reader should be able to obtain more detailed information about each method mentioned by consulting the references.

References

1. Antreich, K.J. and Koblitz, R.K. A New Approach to Design Centering Based on a Multiparameter Yield-Prediction Formula. IEEE Proceedings of 1980 ISCAS, IEEE, 1980.

2. Avriel, Mordecai. *Nonlinear Programming, Analysis and Methods.* Prentice-Hall, Inc., Englewood Cliffs, N. J., 1976.

3. Balaban, Philip. Statistical Simulation of the Fiberguide Repeater and Evaluation of the Error Rate Using Importance Sampling. Proc. of the Int. Symp. on Circ. and Syst., IEEE, 1976, pp. 26-29.

4. Bandler, John W. "Optimization of Design Tolerances Using Nonlinear Programming." *Journal of Optimization Theory and Applications 14*, 14 (1974), 99-114.

5. Bandler, John W., Liu, Peter C. and Chen, James H. K. "Worst Case Network Tolerance Optimization." *IEEE Transactions on Microwave Theory and Techniques MTT-23*, 8 (August 1975), 630-640.

6. Bandler, John W., Liu, Peter C. and Tromp, Herman. "A Nonlinear Programming Approach to Optimal Design Centering, Tolerancing, and Tinin." *IEEE Transactions on Circuits and Systems CAS-23*, 3 (March 1976), 155-165.

7. Bandler, John W. and Abdel-Malek, Hanyl L. "Optimal Centering, Tolerancing, and Yield Determination via Updated Approximations and Cuts." *IEEE Transactions on Circuits and Systems CAS-25*, 10 (October 1978), 853-871.

8. Bandler, J. W. and Rizk, M. R. M. "Optimization of Electric Networks." *Mathematical Programming Study 11* (1979), 1-64.

9. Bohling, Dorothea M. and O'Neill, Lawrence A. "An Interactive Computer Aproach to Tolerance Analysis." *IEEE Trans. on Computers C-19*, 1 (January 1970), 10-16.

10. Brayton, Robert K. and Director, Stephen W. "Computation of Delay Time Sensitivities for Use in Time Domain Optimization." *IEEE Transactions on Circuits and Systems cas-72*, 12 (December 1975), 910-920.

11. Brayton,R. K., Hachtel, G. D. and Director, S. W. Arbitrary Norms for Statistical Design via Linear Programming. Proceedigs of the Intern. Symp. of Circ. and Syst., IEEE, 1978, pp. 161-164.

12. Brayton, Robert K., Director, S. W., Hachtel,Gary D. and Vidigal, L. M. "A New Algorithm for Statistical Circuit Design Based on Quasi-Newton Methods and Function Splitting." *IEEE Transactions on Circuits and Systems CAS-26*, 9 (September 1979), 784-794.

13. Brayton,Robert K., Director,Stephen W. and Hachtel,Gary D. "Yield Maximization and Worst Case Design with Arbitrary Statistical Distributions{." *IEEE Transactions on Circuits and Systems CAS-27*, 9 (September 1980), 756-764.

14. Butler, Edward M. "Realistic Design Using Large-Change Sensitivities and Performance Contours." *IEEE Trans. Circuit Theory CT-18*, 1 (January 1971), 58-65.

15. Cottle.Richard W. and Ferland, Jacques A. "Matrix-Theoretic Criteria for Quasi-Convexity and Pseudo-Convexity of Quadratic Functions." *Linear Algebra and its Applications 5* (1972), 123-136.

16. Crouzeix, Jean-Pierre. "On Second Order Conditions for Quasiconvexity." *Mathematical Programming 18* (1980), 349-352.

17. Cullum, Jane and Brayton, R. K. Some Remarks on the Symmetric Rank One Update. Tech. Rept. RC 6157, IBM, August, 1976.

18. Director, Stephen W. and Rohrer, Ronald A. "Automated Network Design - The Frequency Domain Case." *IEEE Transactions on Circuit Theory ct-16*, 3 (August 1969), 330-337.

19. Director, Stephen W. "Survey of Circuit-Oriented Optimization Techniques." *IEEE Transactions onCircuit Theory CT-18*, 1 (January 1971), 3-10.

20. Director.Stephen W. and Hachtel, Gary D. "The Simplicial Approximation Approach to Design Centering." *IEEE Transactions oc Circuits and Systems CAS-24*, 7 (July 1977), 363-372.

21. Director, Stephen W., Hachtel, Gary D. and Vidigal, Luis Manuel. "Computationally Efficient Yield Estimation Procedures Based on Simplicial Approximation." *IEEE Transactions on Circuits and Systems CAS-25*, 3 (March 1978), 121-130.

22. Elias, N. J. New Statistical Methods For Assigning Device Tolerances. Proc. Intern. Symp. on Circuits and Systems, IEEE, 1975, pp. 329-332.

23. Elias, Norm. Statistical Tolerancing with Data Smoothing the Capitol Circuit Simulator.

24. Jacques A. Ferland. "Maximal Domains of Quasi-Convex and Pseudo-Convexity for Quadratic Functions." *Mathematical Programming 3* (1972), 178-192.

25. Glesner, Manfred and Newmuller, Manfred. Investigation and Approximation of Feasible Regions for Statistical Design of Electrical Networks.

26. Haber, Seymour. "Numerical Evaluation of Multiple Integrals." *SIAM Review 12*, 4 (October 1970), 481-526.

27. Hachtel, Gary, Brayton, Robert K. and Gustavson, Fred G. "The Sparse Tableau Approach to Network Analysis and Design." *IEEE Transactions on Circut Theory CT = 18*, 1 (January 1971), 101-113.

28. Hammersley, J. M. and Handscomb, D. C.. *Monte Carlo Methods.* Methuen, London, 1964.

29. Han,S. P. "A Globally Convergent Method for Nonlinear Programming." *Journal of Optimization Theory and Applications 22*, 3 (July 1977), 297-309.

30. Ilumoka, A.A., Spence, R. and Soin,R.S. The Tolerance Design of Circuits by Statistical Exploration. International Symposium on Circuits and Systems, IEEE, 1980, pp. 878-881.

31. Kahn, Herman. Use of Different Monte Carlo Sampling Techniques. Symposium on Monte Carlo Methods, 1956, pp. 146-190.

32. Karafin, Barry Joel. *The General Component Tolerance Assignment Problem in Electrical Networks.* Ph.D. Th., University of Pennsylvania, 1974.

33. Kjellstrom, Gregor. "Network Otpimization by Random Variation of Component Values." *Ericsson Technics* , 3 (1969), 136-151.

34. Kjellstrom, Gregor. "Optimization of Electrical Networks with Respect to Tolerance Costs." *Ericsson Technics* , 3 (1970), 158-175.

35. Kjellstrom, G. and Taxen, L. On the Efficient Use of Stochastic Optimization in Network Design. Proceedings Int. Symp. on Circuit Theory, 1976, pp. 714-717.

36. Kjellstrom, Gregor and Taxen, Lars. Stochatic Optimization in System Design.

37. Kjellstrom, G., Taxen. L and Blomgren. L-E. Ctpimization Met nods for Statistical Network Oesign. Proceedings int. Symp. on Circuit Theory, 1975, pp. 321-324.

38. Leichner. G. H. *A Statistical Approach to the Tolerance Problem in Oesign.* Ph.D. Th., University of Illinois, 196a

39. Ughtner.Michael R. *Multiple Criterion Optimization and Statistical Design for Electronic Circuits. Ph.D.* Th., Carnegie-Mellon University, 1979.

40. Madsen, Kaj, Schjaer-Jacoben, Hans and Voldby, Jorgen. "Automated Minimax Design of Networks." *IEEE Transactions on Circuits and Systems CAS-22,*10 (October 1975), 791-796.

41. Madsen, K. "An Algorithm for Minimax Solution of Overdetermined Systems of Non-linear Equations." *J. Inst. Maths Applies 16* (1975), 321-32a

42. Madsen, Kaj and Schjaer-Jacobsen, Hans. "Singularities in Minimax Optimization of Networks." *IEEE Transactions on Circuits and Systems CAS-23,7* (July 1976), 456-460.

43. Madsen. Kaj and Schjaer-Jacobsen, Hans. "Linearly Constrained Minimax Optimization." *Mathematical Programming* 74 (1978), 208-223.

44. Madsen,K.andSchaer-Jacobsen,H. New Algorithms for Worst Case Tolerance Optimization, international Symposium on Circuits and Systems, IEEE 1978, pp. 681-685.

45. MaJy.W. and Oirector,S.W. A Dimension Reduction Procedure for the Simpiicial Approximation Approach to Oesign Centering, European Conference on Circuits and Systems, September, 196a

46. Mangasarian,O. L. *Nonlinear Programming.* McGraw-Hill Book Company, 1969.

47. Marios, Beta. "Subdefinite Matrices and Quadratic Forms." *SIAM Journal App. Math. 17,*6 (November 1969), 1215-1223.

48. Martos,Be)a. "Quadratic Programmigwfthaquasiconvex Objective Function." *Operations Research* 79 (1971), 87-97.

49. Ogrodzki,J.,Opalski,LandStybiinski,M. Acceptability Regions for a Class of Linear Networks. Proceedings of the international Symposium on Circuits and Systems, IEEE 1960, pp. 187-190.

50. Pinel, John F. and Roberts, Kenneth A. "Tolerance Assignment in Linear Networks Using Nonlinear Programming." *IEEE Trans, on Circ. Theory* 79, 5 (September 1972), 475-479.

51. Pinel, J. F. and Roberts, K. A. "Tolerance Assignment in Network Design." *IEEE Proc. ISCAS* (1975), 317-320.

52. Pinel, John F. and Singhal, K. Efficient Monte Carlo Computation of Circuit Yield Using Importance Sampling. Proc. ISCAS, IEEE May. 1977, pp. 575-57a

53. Polak, E. and Sangiovanni-Vincentelli, A. On Optimization Algorithms for Engeneering Design Problems with Distributed Constraints, Tolerances and Tuning. Proc. of JACC, ISA, October, 1978, pp. 345-353.

54. Polak, E. and Trahan,R. "Combined Phase I-Phase II Methods of Feasible Directions." *Mathematical Programming 17* (1979), 61-73.

55. Poweil,M.J.D. Rank One Method for Unconstrained Optimization. In *Nonlinear and Integer Programming,* J. Abach, Ed.,North-Holland, Amsterdam, 1970.

56. Powell.M.J.D. A Fast Algo^thm for Nonlinearly Constrained Optimization Calculations. In *#630 Lee. Notes in Math.,* A. Dold and B. Eckmann, Ed.,Springer-Verlag. 1977, pp. 144-157.

57. Powell,MJ.D. Variable Metric Methods for Constrained Optimization. Third International Symposium on Computing Models in Applied Sciences and Engineering, 1977.

58. Powell.M.J.D. "Algorithms for Nonlinear Optimization that use Lagrangian Functions." *Mathmatical Programming 14* (1978), 224-24a

59. S<^jaer-Jacobsen,HansandMadsen.Kai. "Algorithms for Worst-Case Tolerance Optimization." *IEEE Transactions on Circuits and Systems CAS-26,*9 (September 1979), 775-783.

60. Shreider, Y. A.. *The Monte Carlo Method.* Pergammon Press, 1966.

61. Soin, R. S. and Spence, R. "Statistical Design Centering for Electrical Circuits." *Electronics Letters u,* 24 (November 1978), 772-774.

62. Thorbjomsen, A. R. *Computer-Aided Tolerance Oesign.* Ph.D. Th., University of Florida, 1972.

63. VidigaJ, Luis Manuel. *Statistical Design of Electronic Circuits.* Ph.D. Th., Carnegie-Mellon University, August 1980.

64. Sangiovanni-VincenteiU, Alberto. Personal Communication.

65. Weeks, William T., Jimenez, Alberto Jose, Mahoney, Gerald W., Mehta, Deepak, Oassemzadeh, Hassan and Scott, Terence R. "Algorithms for ASTAP- A Network Analysts Program." *IEEE Transactions on Circuit Theory CT-20,*6 (November 1973), 628-634.